## General Instructions

* Do all the documentation in word, write the code and include output snapshots in the word file. Format the document properly.
* Try to solve the issues on your own. Never approach your guide with an open question. Go with a question and two possible answers.
* Keep good folder structure for code, word file and data sets
* Go beyond the class room code, try solving the problem with different code and package options
* The project is not small and it is not easy, be prepared to face it.
* Mention the case study id in all your documents

## Project Steps

1. Objective
   1. Write objective of the Analysis
2. Data
   1. Understand the data and read the data dictionary
   2. Do basic data exploration, Import the data and do some basic data checks, like rows columns etc.,
   3. What are the variable names?
   4. Write the formats of each variable. Are they consistent with data dictionary? Are the variable formats correct?
   5. Explain the variable meanings in simple terms
   6. Are there any missing values in the data?
   7. Divide the dataset into two sets. Training data, consider 80% of the data for training. Testing data, consider remaining 20% of the data for testing.
3. Variable Level Analysis
   1. Go through each and every variable, find out missing values and outliers.
      1. If the variable is continuous then draw box plots, calculate percentiles and Find basic disciple statistics of the variable
      2. If the variable is categorical or discrete, create frequency tables
   2. If you find missing values or outliers try clean the data by imputing
   3. Replace the missing values with mean or median
   4. Replace the missing categories with most probable categories based on target variable
4. Visualizations
   1. Try to give basic summary of the data
   2. For each variable try to identify the interesting patterns and draw the graphs
   3. Create individual graphs and insights for each variable.
   4. Try to use good visualization packages for drawing the graphs
   5. Try to perform bivariate analysis with the target variable and other related variables
5. Document your results
   1. Document your results
   2. Write down your insights